Converse Results for Secrecy Generation over Channels

Himanshu Tyagi[†] Shun Watanabe[‡]

***Abstract*—We revisit the problem of secret key agreement in channel models, where in addition to a noisy, albeit secure channel, the terminals have access to a noiseless public communication channel. We show a strong converse for the secret key capacity in the point-to-point model and give upper bounds for the general case. Underlying our proofs is a recently discovered single-shot converse for secret key rates in multiterminal source models.**

I. Introduction

We revisit the problem of secret key agreement in channel models, where in addition to a noisy, albeit secure broadcast channel, the terminals have access to a noiseless public communication channel. This problem was first investigated by Csiszár and Narayan in [3], and they showed that the maximum rate of a secret key (SK) with asymptotically perfect reliability and secrecy, namely the SK capacity C, is given by[1]

$$C = \max_{\mathrm{P}_{X_0}} \min_{\pi} D\left(\mathrm{P}_{X_{\mathcal{M}}} \middle\| \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi(i)}}\right), \qquad (1)$$

where the maximum is with respect to the input distribution of the sender, and the minimum is with respect to nontrivial partitions of the parties. In this paper, we establish an upper bound on the (ϵ, δ)-SK capacity $C_{\epsilon,\delta}$ for a given reliability ϵ and a given secrecy δ. Specifically, we show that for $\epsilon+\delta < 1$,

$$C_{\epsilon,\delta} \leq \min_{\pi} \max_{\mathrm{P}_{X_0}} D\left(\mathrm{P}_{X_{\mathcal{M}}} \middle\| \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi(i)}}\right),$$

which in turn leads to a strong converse for SK capacity if the minimax equality holds in (1); this is clearly the case for a channel with one output.

Our proof relies on a slight modification of a recent reduction of hypothesis testing to secret key agreement for source models shown in [10], [11]. Specifically, we show that a secret key generation protocol yields an active hypothesis test for distinguishing between two channels [4]. This approach is along the lines of *meta converse* of [7], where a reduction of hypothesis testing to channel coding was used to establish a finite-blocklength converse for the channel coding problem. A similar approach was recently applied to derive the strong converse for the capacity of a degraded wiretap channel with feedback [5].

The rest of the paper is organized as follows: Our main result is given in the next section. Section III and IV contain a review of relevant results in binary hypothesis testing and secret key agreement in source models, respectively. We prove our main result in Section V, followed by a discussion on possible extensions in the final section.

II. Problem Formulation and Main result

Consider a broadcast channel with one input $\mathcal{X}_0$ and m outputs $\mathcal{X}_1, \ldots, \mathcal{X}_m$, specified by a discrete memoryless channel (DMC) $W : \mathcal{X}_0 \to \mathcal{X}_1 \times \cdots \times \mathcal{X}_m$. We study a secrecy generation problem for $(m+1)$ terminals using n transmissions over W where terminal 0 selects the input X_0 for each transmission and terminals $1, ..., m$, respectively, observe the corresponding outputs $X_1, ..., X_m$. For brevity, the set of all terminals is denoted by $\mathcal{M} = \{0, 1, \ldots, m\}$. Between two consecutive transmissions, the terminals communicate with each other interactively over a noiseless public communication channel of unlimited capacity. While the transmissions over the DMC W are secure, the public communication is observed by all the terminals as well as a (passive) eavesdropper. This model was first studied by Csiszár and Narayan in [3]. In the manner of [3], the messages sent over W will be referred to as transmissions and those sent over the public channel will be refereed to as communication.

Formally, assume that at the outset terminal i generates a random variable U_i, $i = 0, 1, \ldots, m$, to be used for (local) randomization; the random variables $U_0, \ldots, U_m$ are mutually independent. The *communication-transmission protocol* can be divided into $n+1$ time slots. In each of the first n time slots, the terminals communicate interactively over the public channel, followed by a transmission over the secure DMC. The protocol ends with a final rounds of interactive public communication in slot $n+1$. Specifically, in time slot t, $1 \leq t \leq n$, the terminals communicate interactively using their respective local randomization $U_0, \ldots, U_m$ and observations up to time slot $t-1$; the overall interactive communication in slot t is denoted by

$$F_t = F_t(U_0, \ldots, U_m, X_1^{t-1}, \ldots, X_m^{t-1}, F^{t-1}).$$

Subsequently, the input $X_{0t} = X_{0t}(F^t, U_0)$ is transmitted by terminal 0, and $X_{1t}, \ldots, X_{mt}$ are observed by terminal

[†] Information Theory and Applications (ITA) Center, University of California, San Diego, La Jolla, CA 92093, USA. Email: htyagi@eng.ucsd.edu

[‡] Department of Information Science and Intelligent Systems, University of Tokushima, Tokushima 770-8506, Japan, and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. Email: shunwata@is.tokushima-u.ac.jp

[1] It follows from a result of [2] that the expression in (1) is an upper bound for the original formula for SK capacity established in [3]. This upper bound was later shown to be tight in [1].

1 through m. Finally, the last round of interactive communication $F_{n+1} = F_{n+1}(U_0, \ldots, U_m, X_1^n, \ldots, X_m^n, F^n)$ is sent over the public channel. For convenience, we denote $\mathbf{F} = (F_1, \ldots, F_{n+1})$.

After the communication-transmission protocol ends, the terminals $0, \ldots, m$, respectively, form estimates $K_0, K_1, \ldots, K_m$ as follows:

$$K_0 = K_0(U_0, \mathbf{F})$$

and

$$K_i = K_i(X_i^n, U_i, \mathbf{F}), \quad i = 1, \ldots, m.$$

A random variable K with range $\mathcal{K}$ constitutes an (ϵ, δ)-SK if the following two conditions are satisfied (*cf.* [3]):

$$\mathrm{P}\left(K_0 = \cdots = K_m = K\right) \geq 1 - \epsilon, \tag{2}$$

$$\|\mathrm{P}_{K\mathbf{F}} - \mathrm{P}_{\mathsf{unif}} \times \mathrm{P}_{\mathbf{F}}\|_1 \leq \delta, \tag{3}$$

where $\mathrm{P}_{\mathsf{unit}}$ is the uniform distribution on $\mathcal{K}$, and where $\|\mathrm{P} - \mathrm{Q}\|_1$ denote the (normalized) variation distance between P and Q given by

$$\|\mathrm{P} - \mathrm{Q}\|_1 = \frac{1}{2}\sum_x |\mathrm{P}\,(x) - \mathrm{Q}(x)|.$$

The first condition above represents the *reliability* of the SK and the second guarantees its *secrecy*.

Definition 1. A rate $R > 0$ is (ϵ, δ)*-achievable* if there exists local randomization $U_0, \ldots, U_m$, communication-transmission protocol $\mathbf{F}$, and (ϵ, δ)-SK K with

$$\frac{1}{n}\log|\mathcal{K}| \geq R$$

for all sufficiently large n. The supremum of all (ϵ, δ)-achievable rates is called the (ϵ, δ)*-SK capacity*, denoted by $C_{\epsilon,\delta}$.

Our main result is an upper bound on $C_{\epsilon,\delta}$.

Theorem 1. *For $0 \leq \epsilon, \delta$ with $\epsilon + \delta < 1$, the (ϵ, δ)-SK capacity is bounded above as*

$$C_{\epsilon,\delta} \leq \min_{\pi} \max_{\mathrm{P}_{X_0}} D\left(\mathrm{P}_{X_\mathcal{M}} \middle\| \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi(i)}}\right),$$

where min *is taken over all nontrivial partitions, and where $\mathrm{P}_{X_\mathcal{M}}$ the joint distribution induced from input distribution P_{X_0} and channel W.*

Clearly, when there is only one receiver ($m = 1$) the strong converse for SK capacity follows.

Corollary 2. *For SK generation with one receiver, the (ϵ, δ)-SK capacity is given by*

$$C_{\epsilon,\delta} = \max_{\mathrm{P}_{X_0}} I(X_0 \wedge X_1)$$

for every $0 < \epsilon < 1 - \delta$.

Note that $1 - \delta \leq \epsilon < 1$ the (ϵ, δ)-SK capacity is unbounded. Therefore, the corollary above characterizes $C_{\epsilon,\delta}$ for all $0 < \epsilon, \delta < 1$.

III. Hypothesis testing

Consider a simple binary hypothesis testing problem with null hypothesis P and alternative hypothesis Q, where P and Q are distributions on the same alphabet $\mathcal{X}$. Upon observing a value $x \in \mathcal{X}$, the observer needs to decide if the value was generated by the distribution P or the distribution Q. To this end, the observer applies a stochastic test T, which is a conditional distribution on $\{0, 1\}$ given an observation $x \in \mathcal{X}$. When $x \in \mathcal{X}$ is observed, the test T chooses the null hypothesis with probability $\mathrm{T}(0|x)$ and the alternative hypothesis with probability $T(1|x) = 1 - T(0|x)$. For $0 \leq \epsilon < 1$, denote by $\beta_\epsilon(\mathrm{P}, \mathrm{Q})$ the infimum of the probability of error of type II given that the probability of error of type I is less than ϵ, i.e.,

$$\beta_\epsilon(\mathrm{P}, \mathrm{Q}) := \inf_{\mathrm{T}\,:\,\mathrm{P}[\mathrm{T}] \geq 1-\epsilon} \mathrm{Q}[\mathrm{T}],$$

where

$$\begin{aligned}\mathrm{P}[\mathrm{T}] &= \sum_x \mathrm{P}(x)\mathrm{T}(0|x),\\ \mathrm{Q}[\mathrm{T}] &= \sum_x \mathrm{Q}(x)\mathrm{T}(0|x).\end{aligned}$$

The following result credited to Stein characterizes the optimum exponent of $\beta_\epsilon(\mathrm{P}^n, \mathrm{Q}^n)$ where $\mathrm{P}^n = \mathrm{P} \times \ldots \times \mathrm{P}$ and $\mathrm{Q}^n = \mathrm{Q} \times \ldots \times \mathrm{Q}$.

Lemma 3. (*cf.* [6, Theorem 3.3]) *For every $0 < \epsilon < 1$, we have*

$$\lim_{n\to\infty} -\frac{1}{n}\log\beta_\epsilon(\mathrm{P}^n, \mathrm{Q}^n) = D(\mathrm{P}\|\mathrm{Q}),$$

where $D(\mathrm{P}\|\mathrm{Q})$ is the Kullback-Leibler divergence given by

$$D(\mathrm{P}\|\mathrm{Q}) = \sum_{x\in\mathcal{X}} \mathrm{P}(x)\log\frac{\mathrm{P}(x)}{\mathrm{Q}(x)},$$

with the convention $0\log(0/0) = 0$.

Next, we review a problem of active hypothesis testing where the distribution at each instance is determined by a prior action. Specifically, given two DMCs $W : \mathcal{X} \to \mathcal{Y}$ and $V : \mathcal{X} \to \mathcal{Y}$, we seek to design a transmission-feedback scheme such that by observing the channel inputs, channel outputs, and feedback we can determine if the underlying channel is W or V. Formally, an n-length active hypothesis test consist of (possibly randomized) encoder mappings $e_t : \mathcal{F}^t \to \mathcal{X}$, $1 \leq t \leq n$, feedback mappings $f_t : \mathcal{Y}^t \to \mathcal{F}$, $0 \leq t \leq n-1$, and a conditional distribution T on $\{0, 1\}$ given $X^n, Y^n, \mathbf{F}$. On observing $X^n, Y^n, \mathbf{F}$, we detect the null hypothesis W with probability $T(0|X^n, Y^n, \mathbf{F})$ and alternative hypothesis V with probability $T(1|X^n, Y^n, \mathbf{F})$. Analogous to $\beta_\epsilon(\mathrm{P}, \mathrm{Q})$, the quantity $\beta_\epsilon(W, V, n)$, for $0 \leq \epsilon < 1$, is the infimum of the probability of error of type II over all n length active hypothesis tests for null hypothesis W and alternative hypothesis V such that the probability of error of type I is no more than ϵ.

The following analogue of Stein's lemma for active hypothesis testing was established in [4] (see, also, [8]).

Theorem 4 ([4]). *For $0 < \epsilon < 1$,*

$$\lim_n -\frac{1}{n}\log\beta_\epsilon(W,V,n) = \max_{\mathrm{P}_X} D\big(W\big\|V \,\big|\, \mathrm{P}_X\big) = \max_x D\big(W_x\big\|V_x\big),$$

where W_x and V_x, respectively, denote the xth row of W and V.

Remarkably, the exponent above is achieved without any feedback, *i.e.*, while feedback is available, it does not help to improve the asymptotic exponent of $\beta_\epsilon(W,V,n)$. Furthermore, the proof in [4] remains valid even when the feedback f_t is replaced by a noiseless interactive communication.

IV. Converse for Source Model

In this section, we review multiparty SK agreement where parties observing random variables $X_0, \ldots, X_m$ communicate interactively over a public channel to agree on a SK that is concealed from an eavesdropper with access to the communication.

Formally, the communication is sent over r rounds of interaction. In the jth round of communication, $1 \leq j \leq r$, the ith party sends F_{ij}, which is a function of its observation X_i, local randomness U_i, and the previously observed communication. At the end of the protocol, the ith party computes an estimate $K_i = K_i(X_i, \mathbf{F})$ of the SK. A random variable K with range $\mathcal{K}$ constitute an (ϵ,δ)-SK if conditions (2) and (3) are satisfied.

The following upper bound on the length of the key taken by an (ϵ,δ)-SK K was shown in [10], [11]:

$$\log|\mathcal{K}| \leq \min_\pi \frac{1}{|\pi|-1}\left[-\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{X_\mathcal{M}}, \mathrm{Q}^\pi_{X_\mathcal{M}}) + |\pi|\log\frac{1}{\eta}\right],$$

for all $0 < \eta < 1-\epsilon-\delta$, all nontrivial partitions π, and all $\mathrm{Q}^\pi_{X_\mathcal{M}} = \prod_{i=1}^{|\pi|}\mathrm{Q}_{X_{\pi(i)}}$. Underlying the proof of this bound is an intermediate reduction argument in [10, Lemma 1] that relates SK agreement to hypothesis testing. We recall this result below.

Theorem 5 ([10], [11]). *For every $0 < \eta < 1-\epsilon-\delta$ and every $\mathrm{Q}^\pi_{K_\mathcal{M}\mathbf{F}}$ satisfying*

$$\mathrm{Q}^\pi_{K_\mathcal{M}|\mathbf{F}} = \prod_{i=1}^{|\pi|}\mathrm{Q}_{K_{\pi(i)}|\mathbf{F}},$$

the length of (ϵ,δ)-SK satisfies

$$\log|\mathcal{K}| \leq \frac{1}{|\pi|-1}\left[-\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{K_\mathcal{M}\mathbf{F}}, \mathrm{Q}^\pi_{K_\mathcal{M}\mathbf{F}}) + |\pi|\log\frac{1}{\eta}\right].$$

V. Proof of main result

We present a converse result that applies for every fixed n and is asymptotically yields Theorem 1.

Theorem 6. *For $0 \leq \epsilon, \delta, \epsilon+\delta < 1$, given a communication-transmission protocol generating (ϵ,δ)-SK, we have*

$$\log|\mathcal{K}| \leq \frac{1}{|\pi|-1}\left[-\log\beta_{\epsilon+\delta+\eta}(W, V^\pi, n) + |\pi|\log\frac{1}{\eta}\right],$$

for all $0 < \eta < 1-\epsilon-\delta$, all non-trivial partition π, and all channels $V^\pi : \mathcal{X}_0 \to \mathcal{X}_1 \times \cdots \times \mathcal{X}_m$ satisfying the following:

$$V^\pi(x_1,\ldots,x_m|x_0) = V_1(x_{\pi(1)\setminus\{0\}}|x_0)\prod_{i=2}^{|\pi|}V_i(x_{\pi(i)}), \quad (4)$$

where $\pi(1) \subset \mathcal{M}$ is the subset that contains 0 and $V_i(\cdot)$ is a channel that does not depend on the input x_0.

Proof of Theorem 1. Let $\mathrm{P}^*_{X_0}$ be a maximizer of

$$\max_{\mathrm{P}_{X_0}} D\left(\mathrm{P}_{X_\mathcal{M}}\middle\|\prod_{i=1}^{|\pi|}\mathrm{P}_{X_{\pi(i)}}\right),$$

and let $\mathrm{P}^*_{X_\mathcal{M}}$ be the corresponding distribution of $X_\mathcal{M}$. Theorem 1 follows form Theorems 6 and 4 upon setting $V_i = \mathrm{P}_{X_{\pi(i)}}$ for $i = 2,\ldots,|\pi|$and $V_1(x_{\pi(1)\setminus\{0\}}|x_0) = W(x_{\pi(1)\setminus\{0\}}|x_0)$. □

Finally, we prove Theorem 6; the following observation will be used.

Lemma 7. *For a channel V^π of the form (4) and a given communication-transmission protocol, the induced distribution $\mathrm{Q}^\pi_{K_\mathcal{M}\mathbf{F}}$ satisfies the factorization condition*

$$\mathrm{Q}^\pi_{K_\mathcal{M}|\mathbf{F}} = \prod_{i=1}^{|\pi|}\mathrm{Q}_{X_{\pi(i)}|\mathbf{F}}.$$

Proof of Lemma 7. The proof entails a repeated application of the fact that conditionally independent random variables remain so when conditioned additionally on an interactive communication ($cf.$ [9]) and is completed by induction. Specifically, note first that $\mathrm{Q}_{U_0\cdots U_m|F_0}$ satisfies the factorization condition since the local randomness are independent and F_0 is an interactive communication. Under the induction hypothesis

$$\mathrm{Q}^\pi_{U_\mathcal{M}X_\mathcal{M}|F^{t-1}} = \prod_{i=1}^{|\pi|}Q_{U_{\pi(i)}X^{t-1}_{\pi(i)}|F^{t-1}},$$

we get

$$\begin{aligned} &I(U_{\pi(i)}X^t_{\pi(i)} \wedge (U_{\pi(j)}, X^t_{\pi(j)} : 1 \leq j \neq i \leq |\pi|)|F^{t-1}) \\ &= I(U_{\pi(i)}X^{t-1}_{\pi(i)} \wedge (U_{\pi(j)}, X^{t-1}_{\pi(j)} : 1 \leq j \neq i \leq |\pi|)|F^{t-1}) \\ &= 0, \end{aligned}$$

where the first equality follows since $(X_{\pi(j),t} : 2 \leq j \leq |\pi|)$ are the outputs of V_js, $X_{0,t}$ is a function of (U_0, F^{t-1}), and $X_{\pi(1)\setminus\{0\},t}$ is the outputs of V_1. □

Proof of Theorem 6. Given a communication-transmission protocol, let $Q^\pi_{K_\mathcal{M}\mathbf{F}}$ be the distribution induced by the protocol when the underlying channel is a V^π of the form (4). Then, by Lemma 7, $Q^\pi_{K_\mathcal{M}\mathbf{F}}$ satisfies the condition of

Theorem 5 and the following bound holds:

$$\begin{aligned}&\log|\mathcal{K}|\\&\leq \frac{1}{|\pi|-1}\left[-\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}},\mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}})+|\pi|\log\frac{1}{\eta}\right].\end{aligned}$$

Note that a test for the simple binary hypothesis testing problem for $\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}}$ and $\mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}}$ along with the communication-transmission protocol constitutes an active hypothesis test for W and V. Therefore,

$$\begin{aligned}&-\log\beta_{\epsilon+\delta+\eta}(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}},\mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}})\\&\quad\leq-\log\beta_{\epsilon+\delta+\eta}(W,V^{\pi}),\end{aligned}$$

which completes the proof. □

VI. Discussion

Our approach yields a strong converse for SK capacity in Corollary 2 for the case $m = 1$. In order to extend this result to a general m, an extension of the result of [4] to "composite" active hypothesis testing with finitely many alternative channels is needed. In particular, it needs to be shown that feedback does not improve the exponent probability of error of type II even for the composite case. The validity of this statement remains unclear.

References

[1] C. Chan and L. Zheng, "Mutual dependence for secret key agreement," *Proc. Annual Conference on Information Sciences and Systems (CISS)*, 2010.

[2] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, December 2004.

[3] ——, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2437–2452, June 2008.

[4] M. Hayashi, "Discrimination of two channels by adaptive methods and its application to quantum system," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3807–3820, Aug 2009.

[5] M. Hayashi, H. Tyagi, and S. Watanabe, "Strong converse for a degraded wiretap channel via active hypothesis testing," *Proc. Conference on Communication, Control, and Computing (Allerton)*, 2014.

[6] S. Kullback, *Information Theory and Statistics*. Dover Publications, 1968.

[7] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[8] Y. Polyanskiy and S. Verdú, "Arimoto channel coding converse and Rényi divergence," *Proc. Conference on Communication, Control, and Computing (Allerton)*, pp. 1327–1333, 2010.

[9] H. Tyagi and P. Narayan, "How many queries will resolve common randomness?" *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5363–5378, September 2013.

[10] H. Tyagi and S. Watanabe, "A bound for multiparty secret key agreement and implications for a problem of secure computing," in *Proc. EUROCRYPT*, 2014, pp. 369–386.

[11] ——, "Converses for secret key agreement and secure computing," *CoRR*, vol. abs/1404.5715, 2014.